

**UNITED STATES DEPARTMENT OF COMMERCE****United States Patent and Trademark Office**

Address: COMMISSIONER OF PATENTS AND TRADEMARKS  
Washington, D.C. 20231

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.
-----------------	-------------	----------------------	---------------------

09/392,844 09/09/99 AUGUST

K AUGUST25-1-7

TM02/0829  
FAY SHARPE FAGAN MINNICH & MCKEE LLP  
1100 SUPERIOR AVENUE SEVENTH FLOOR  
CLEVELAND OH 44114

EXAMINER

LERNER, M

ART UNIT

PAPER NUMBER

2641

DATE MAILED:

08/29/01

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner of Patents and Trademarks

# Office Action Summary

Application No.

09/392,844

Applicant(s)

AUGUST ET AL.

Examiner

Martin Lerner

Art Unit

2641

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

## Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

## Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is FINAL. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

## Disposition of Claims

- 4) ☒ Claim(s) 1 to 19 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1 to 19 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

## Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) ☐ The proposed drawing correction filed on \_\_\_\_ is: a) ☐ approved b) ☐ disapproved by the Examiner.
- If approved, corrected drawings are required in reply to this Office action.
- 12) ☐ The oath or declaration is objected to by the Examiner.

## Priority under 35 U.S.C. §§ 119 and 120

- 13) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
- a) ☐ The translation of the foreign language provisional application has been received.
- 15) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

## Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892)
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) 5.
- 4) ☐ Interview Summary (PTO-413) Paper No(s). \_\_\_\_.
- 5) ☐ Notice of Informal Patent Application (PTO-152)
- 6) ☐ Other: \_\_\_\_\_.

## **DETAILED ACTION**

### ***Claim Objections***

1. Claims 1 to 15 are objected to because of the following informalities:

In claim 1, line 5, "a user" should be "the user" since the preamble already recites "a user." Appropriate correction is required.

### ***Claim Rejections - 35 USC § 102***

2. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

3. Claims 1, 5 to 8, 10 and 15 are rejected under 35 U.S.C. 102(e) as being anticipated by *Adams, Jr. et al.*

Regarding independent claim 1, *Adams, Jr. et al.* discloses an interactive language learning system, comprising:

“a first module configured to convert input text to audible speech in a selected language, the audible speech being patterned after a model” – visually displayed text segments will be accompanied by simultaneously rendered audio presentation of the text segments from audio output 14; the audio text segments may be synthesized in real time by the executive program (column 7, lines 9 to 15: Figure 2); the system represents a computer companion to share the task of reading or language learning (column 2, lines 42 to 43; column 1, lines 27 to 29); implicitly, text-to-speech synthesis is based upon phoneme models of speech;

“a user interface configured to receive utterances spoken by a user in response to a prompt to replicate the audible speech” – the positional pacer 17, with input from the local text position management databases 22, may be implemented to continuously prompt the student along the text, identifying the word that is to be pronounced by the student (column 7, lines 20 to 25: Figure 2);

“a second module configured to recognize the utterances and provide feedback to the user as to a precision at which the user replicates the audible speech in the selected language based on a comparison of the utterances to one of the audible speech and the model” – speech recognition engine 1 is provided with information from the learner population specific acoustic model 3 to facilitate recognition of the audio input (column 5, lines 15 to 32: Figure 2); as the student progresses through the lessons in the application, the executive

program 6 regularly updates the reading level information database with a revised estimate of the student's competency based upon his or her performance (column 8, lined 27 to 43: Figure 2); if the response is not correct, the executive program may access the feedback message database and generate appropriate feedback; if a correct response is received, optional positive feedback may be generated (column 9, lines 40 to 47: Figures 4A and 4B: Steps 208, 211).

Regarding claim 5, *Adams, Jr. et al.* discloses acoustic models based on phonemes (column 5, lines 15 to 32; column 5, lines 57 to 65).

Regarding claim 6, *Adams, Jr. et al.* discloses alternate phrase and pronunciation database 4 and text power set database 12 are additionally provided to enhance recognition of the responses uttered by the student user; both of these latter databases 4, 12, are based upon the text of the story to be read; alternate phrase and pronunciation database provides alternate correct pronunciations (column 5, line 33 to column 6, line 19: Figure 2).

Regarding claim 7, *Adams, Jr. et al.* discloses a story text database 10 which includes the text of the story to be read for each lesson level (column 8, lines 44 to 61: Figures 2 and 3).

Regarding claim 8, *Adams, Jr. et al.* discloses that the student may load a lesson program from a CD-ROM or diskette, or download same from a dedicated network or the internet (column 3, lines 52 to 56).

Regarding claim 10, *Adams, Jr. et al.* discloses that throughout the lesson, audio inputs from both the student and the computer instructor, along with text for each utterance, are stored at the session database for replay and resumption; to review the lesson, the student would click on a “playback icon” (column 7, line 45 to column 8, line 10: Figure 2: 17).

Regarding claim 15, *Adams, Jr. et al.* discloses an alternate phrase and pronunciation database 4, which is a phonemic representation of different ways to pronounce words in the currently active vocabulary (i.e. in the text which is being read at any given time) (column 4, lines 52 to 56: Figure 2).

### ***Claim Rejections - 35 USC § 103***

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. Claims 2 to 4, 9 and 16 to 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Adams, Jr. et al.* in view of *Henton*.

Concerning independent claim 16, *Adams, Jr. et al.* discloses an interactive language learning system, comprising:

“a first module configured to convert input text to audible speech in a selected language, the audible speech indicative of a model” – visually displayed text segments will be accompanied by simultaneously rendered audio presentation of the text segments from audio output 14; the audio text segments may be synthesized in real time by the executive program (column 7, lines 9 to 15: Figure 2); the system represents a computer companion to share the task of reading or language learning (column 2, lines 42 to 43; column 1, lines 27 to 29); implicitly, text-to-speech synthesis is based upon phoneme models of speech;

“a user interface configured to receive utterances spoken by a user in response to a prompt to replicate the audible speech” – the positional pacer 17, with input from the local text position management databases 22, may be implemented to continuously prompt the student along the text, identifying the word that is to be pronounced by the student (column 7, lines 20 to 25: Figure 2);

“a third module configured to recognize the utterances and provide feedback to the user as to a precision at which the user replicates the audible speech in the selected language based on a comparison of the utterances to one of the audible speech and the model” – speech recognition engine 1 is provided with information from the learner population specific acoustic model 3 to facilitate recognition of the audio input (column 5, lines 15 to 32: Figure 2); as the student progresses through the lessons in the application, the executive

program 6 regularly updates the reading level information database with a revised estimate of the student's competency based upon his or her performance (column 8, lined 27 to 43: Figure 2); if the response is not correct, the executive program may access the feedback message database and generate appropriate feedback; if a correct response is received, optional positive feedback may be generated (column 9, lines 40 to 47: Figures 4A and 4B: Steps 208, 211).

*Adams, Jr. et al.* discloses visually displayed text segments as the audio text segments are synthesized in real time, an audio prompt ("follow the bouncing ball"), highlighting and color differentiating the text, and a picture prompt database (column 7, lines 9 to 30), but omits:

"a second module synchronized to the first module, the second module producing an animated image of a human face and head pronouncing the audible speech."

However, *Henton* teaches a method and apparatus for synthetic speech in facial animation, suggesting that it is well known to synchronize facial imaging with synthetic speech for the purpose of instructing the user. (Column 3, Lines 33 to 39: Figure 3) It would have been obvious to one of ordinary skill in the art to include a facial animation module in *Adams, Jr. et al.* that synchronizes imaging with synthetic speech as taught by *Henton* because it is well known to utilize facial animation synchronized with synthetic speech in various applications including user instruction.



Concerning independent claim 17, *Adams, Jr. et al.* discloses an interactive language learning method, comprising:

“converting input text data to audible speech data” – visually displayed text segments will be accompanied by simultaneously rendered audio presentation of the text segments from audio output 14; the audio text segments may be synthesized in real time by the executive program (column 7, lines 9 to 15: Figure 2);

“generating audible speech comprising phonemes based on the audible speech data” – implicitly, speech synthesized from text is based on phonetic sound segments;

“outputting the audible speech through an audio output device” – audio is output via speaker 14 (column 3, lines 32 to 36: Figure 2);

“prompting the user to replicate the audible speech” – a positional pacer 17 (“follow the bouncing ball”) may be implemented to continuously prompt the student along the text, identifying the word that is to be pronounced by the student (column 7, lines 20 to 27);

“recognizing utterances generated by the user in response to the prompting” – speech recognition engine 1 is provided with information to facilitate recognition of the audio input (column 5, lines 15 to 19: Figure 2);

“comparing the audible speech to the utterances” – the executive program evaluates the response and determines whether a correct or incorrect

response is received (column 9, lines 35 to 51: Figures 4A and 4B: Steps 208 and 211);

“providing feedback to the user based on the comparison” -- as the student progresses through the lessons in the application, the executive program 6 regularly updates the reading level information database with a revised estimate of the student’s competency based upon his or her performance (column 8, lines 27 to 43: Figure 2); the executive program may access the feedback message database and generate appropriate feedback (column 9, lines 40 to 47: Figures 4A and 4B: Steps 208, 211).

*Adams, Jr. et al.* discloses visually displayed text segments as the audio text segments are synthesized in real time, an audio prompt (“follow the bouncing ball”), highlighting and color differentiating the text, and a picture prompt database (column 7, lines 9 to 30), but omits “generating an animated image of a face and head pronouncing the audible speech” and “synchronizing the audible speech and the video image.” However, *Henton* teaches a method and apparatus for synthetic speech in facial animation, suggesting that it is well known to synchronize facial imaging with synthetic speech for the purpose of instructing the user. (Column 3, Lines 33 to 39: Figure 3) It would have been obvious to one of ordinary skill in the art to include a facial animation module in *Adams, Jr. et al.* that synchronizes imaging with synthetic speech as taught by *Henton* because it is well known to utilize facial animation

synchronized with synthetic speech in various applications including user instruction.

Concerning claim 2, similar consideration apply.

Concerning claim 3, *Henton* teaches a face and head which is a “transparent” line drawing (Figure 3).

Concerning claim 4, *Adams, Jr. et al.* must implicitly include at least a volume control for speaker 14.

Concerning claim 18, *Adams, Jr. et al.* discloses that the student may load a lesson program from a CD-ROM or diskette, or download same from a dedicated network or the internet (column 3, lines 52 to 56).

Concerning claim 19, *Adams, Jr. et al.* discloses that throughout the lesson, audio inputs from both the student and the computer instructor, along with text for each utterance, are stored at the session database for replay and resumption; to review the lesson, the student would click on a “playback icon” (column 7, line 45 to column 8, line 10: Figure 2: 17).

6. Claims 9 and 11 to 14 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Adams, Jr. et al.* in view of *Mostow et al.*

Concerning claim 9, *Adams, Jr. et al.* includes databases containing various sorts of information (column 4, line 1 to column 5, line 4), but does not specifically disclose dictionary files. However, *Mostow et al.* teaches a related

reading and pronunciation tutor involving speech recognition, suggesting a knowledge base 24 with a lexicon of word pronunciations and definitions of selected words from the text that are believed to be unfamiliar to the reader. (Column 8, Lines 43 to 61: Figure 1) It would have been obvious to include dictionary files in the interactive language instruction system of *Adams, Jr. et al.* as suggested by *Mostow et al.* for the purpose of providing the student with information on words from a foreign language.

Concerning claims 11 to 13, *Adams, Jr. et al.* discloses an alternate phrase and pronunciation database 4 and text power set database 12 for refining the speech recognition process, but omits tables storing mapping data between word subgroups and vocabulary words, between words and vocabulary words, and between words and examples of parts of speech. However, *Mostow et al.* teaches a related reading and pronunciation tutor where an automatic enhancement function includes a heuristic algorithm using tables. Lookup of information in tables identifies sets of words that rhyme with one another, words that look alike, start or end the same etc., by constructing a key for each word that says what set is that word's equivalence class. The word may also be decomposed into its root word and affixes, which implicitly involves identification of the word's part of speech (Column 9, Line 52 to Column 10, Line 33) It would have been obvious to one of ordinary skill in the art to include tables of related words as taught by *Mostow et al.* in the

Art Unit: 2641

interactive language instruction system of *Adams, Jr. et al.* for the purpose of inferring the pronunciation of words not found in a dictionary.

Concerning claim 14, *Adams, Jr. et al.* omits tables of pronunciation, but *Mostow et al.* teaches that the tutoring function takes account of phrase boundaries as indicated by commas and certain other punctuation for the purpose of more accurately aligning recognition results against the text. It would have been obvious to one of ordinary skill in the art to include a table of punctuation indicating phrase boundaries in the interactive language instruction system of *Adams, Jr. et al.* for the purpose of more accurately aligning recognition results against the text as taught by *Mostow et al.*

### **Conclusion**

7. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Poggio et al., Trower, II et al., Waters et al., Sabourin, Hata et al., Tolin et al. and Shpiro et al. disclose related art.

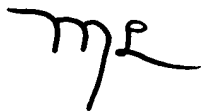
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (703) 308-9064. The examiner can normally be reached on 9:30 AM to 6:00 PM Monday to Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, William Korzuch can be reached on (703) 305-6137.

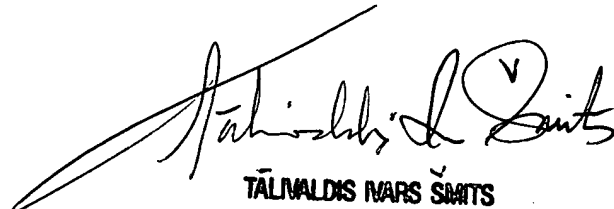
Art Unit: 2641

The fax phone numbers for the organization where this application or proceeding is assigned are (703) 305-9508 for regular communications and (703) 305-9508 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-4700.



ml  
August 23, 2001



TĀLMĀLDIS NARS ŠMITS  
PRIMARY EXAMINER